RECEIVED
CENTRAL FAX CENTER
SEP 18 2006

REMARKS

Claims 1-53 are pending. Claims 1, 9, 17-23, 31, 35-40, 46, 48, 52, and 53 have been amended. Claims 32-34, 45, and 49-51 have been canceled. Claims 1-31, 35-44, 46-48, 52, and 53 remain in the application. The specification has been

5   amended. No new matter has been entered.

The specification stands objected for informalities. The specification has been amended as suggested. The specification has also been amended to update references to commonly-assigned patent applications, which have issued since the time of the filing of this application. No new matter has been entered. Withdrawal

10   of the objection is requested.

Claims 17 and 52 stand objected for improper dependent form. Claims 17 and 52 have been rewritten in independent form. Withdrawal of the objection is requested.

Claims 1-53 stand rejected under 35 U.S.C. § 101 as directed to non-

15   statutory subject matter. Applicant traverses the rejection.

The subject matter of a properly construed claim is defined by the terms that limit its scope. MPEP 2106. Language that _suggests or makes optional_ but does not require steps to be performed or does not limit a claim to a particular structure does not limit the scope of a claim or claim limitation (emphasis

20   original). _Id._ "While it is appropriate to use the specification to determine what applicant intends a term to mean, a positive limitation from the specification _cannot be read into a claim that does not itself impose that limitation_ (emphasis added)." _Id._

On page 9, lines 20-22, the specification provides that "various

25   implementations of the source code and object and byte codes _can_ be held on a computer-readable storage medium _or_ embodied on a transmission medium in a carrier wave. (emphasis added)" This language does not, and is not intended to, "require steps to be performed or . . . limit a claim to a particular structure." MPEP 2106. The language merely "suggests or makes optional but does not

30   require" two possible structures upon which the source code and object and byte

- 17 -

codes can be held. Other structures are possible. Moreover, independent system Claim 1, as amended, recites a scoring module, clustering module, and threshold module. No limitations to "a computer-readable storage medium or embodied on a transmission medium in a carrier wave" are recited. The same observation can

5   be made about independent method Claim 9, which, as amended, recites determining a score, forming logically-grouped clusters, dynamically determining a threshold, and reassigning documents. No steps restricted to "a computer-readable storage medium or embodied on a transmission medium in a carrier wave" are recited. The same observation can be made about independent system

10   Claim 18, which, as amended, recites a scoring module that comprises a frequency module, concept weight module, structural weight module, corpus weight module, and scoring evaluation module; and a clustering module that comprises a cluster seed module and threshold module. No limitations to "a computer-readable storage medium or embodied on a transmission medium in a

15   carrier wave" are recited. Finally, the same observation can be made about independent method Claim 35, which, as amended, recites scoring a document, comprising determining a frequency, analyzing a concept weight, analyzing a structural weight, analyzing a corpus weight, and evaluating a score; and grouping documents, comprising identifying candidate seed documents, assigning each

20   non-seed document, dynamically determining a threshold, and reassigning the documents. No steps restricted to "a computer-readable storage medium or embodied on a transmission medium in a carrier wave" are recited. Reading such limitations into independent Claims 1, 9, 18, 35, and the claims dependent thereon, Claims 2-8, Claims 10-16, Claims 19-31, and Claims 36-44 and 46-48,

25   respectively, and independent Claim 53 is improper.

Claims 17 and 52 have been rewritten in independent form to define computer-readable storage mediums holding code with positive recitations of structural elements. A claimed computer-readable medium encoded with a computer program is a computer element, which defines structural and functional

30   interrelationships between the computer program and the rest of the computer that

- 18 -

permit the computer program's functionality to be realized, and <u>is thus statutory</u>
(emphasis added). MPEP 2106.01(I). Claims 17 and 52 are "computer-readable
medium encoded with a computer program," which are statutory.

5   Accordingly, a *prima facie* case of non-statutory subject matter has not
been shown. Withdrawal of the rejection of remaining Claims 1-31, 35-44, 46-48,
52, and 53 under 35 U.S.C. § 101 is requested.

Claims 19-23, 31, 36-40, and 48 stand rejected under 35 U.S.C. § 112,
second paragraph, for indefiniteness. Claims 19-23, 31, 36-40, and 48 have been
amended to remove the word "substantially." No new matter has been entered.

10   Withdrawal of the rejection under 35 U.S.C. § 112, second paragraph, is
requested.

Claims 1-18 stand rejected under 35 U.S.C. § 102(a) as being anticipated
by International Application Publication No. WO 03/060,766, to Lindh et al.
("Lindh"). Applicant traverses the rejection.

15   Under 35 U.S.C. § 102(b), a claim is anticipated only if each and every
element as set forth in the claim is found, either expressly or inherently described,
in a single prior art reference. *See, Crown Ops. Int'L, Ltd. v. Solutia Inc.*, 289 F.3d
1367 (Fed. Cir. 2002). If each and every element of a claim is not described in the
reference, the claim has not been anticipated.

20   Claim 1 has been amended to further clarify the inventive subject matter
and distinguish over Lindh. No new matter has been entered. Amended Claim 1
now recites a clustering module forming clusters of the documents by evaluating
the score for the at least one concept of each document for a best fit to the clusters
and assigning each document to the cluster with the best fit. Support can be found
25   in the specification on page 9, lines 6-11 and page 18, line 22-page 20, line 3.
Claim 1 further recites a threshold module dynamically determining a threshold
for each cluster based on similarities between the documents grouped into the
cluster and a center of the cluster, and reassigning those documents having
similarities outside the threshold. Support can be found in the specification on
30   page 20, line 4-page 21, line 18.

- 19 -

Claim 9 has also been amended to further clarify the inventive subject matter and distinguish over Lindh. No new matter has been entered. Amended Claim 9 now recites forming logically-grouped clusters of the documents by evaluating the score for the at least one concept of each document for a best fit to

5   the clusters and assigning each document to the cluster with the best fit. Support can be found in the specification on page 18, line 22-page 20, line 3. Claim 9 further recites dynamically determining a threshold for each cluster based on similarities between the documents grouped into the cluster and a center of the cluster; and reassigning those documents having similarities outside the threshold.

10   Support can be found in the specification on page 20, line 4-page 21, line 18.

Claim 17 has been rewritten in independent form to define computer-readable storage mediums holding code with positive recitations of structural elements. Claim 17 similarly distinguishes over Lindh. Claim 17 recites code for forming logically-grouped clusters of the documents by evaluating the score for

15   the at least one concept of each document for a best fit to the clusters and assigning each document to the cluster with the best fit. Support can be found in the specification on page 9, lines 17-24 and page 18, line 22-page 20, line 3. Claim 17 further recites code for dynamically determining a threshold for each cluster based on similarities between the documents grouped into the cluster and a

20   center of the cluster; and code for reassigning those documents having similarities outside the threshold. Support can be found in the specification on page 9, lines 17-24 and page 20, line 4-page 21, line 18

Claim 18 has also been amended to further clarify the inventive subject matter and distinguish over Lindh. No new matter has been entered. Amended

25   Claim 18 now recites a scoring module scoring a document in an electronically-stored document set and a clustering module grouping the documents by score into a plurality of clusters. Support can be found in the specification on page 8, line 26-page 9, line 11. Claim 18 further recites the clustering module comprises a cluster seed module identifying candidate seed documents, which are each

30   assigned as a seed document into a cluster with a center most similar to the seed

- 20 -

document, and assigning each non-seed document to the cluster with the best fit. Support can be found in the specification on page 8, line 26-page 9, line 11 and page 18, line 22-page 20, line 3. Claim 18 further recites a threshold module dynamically determining a threshold for each cluster based on similarities

5   between the documents in each cluster and the cluster center, and reassigning the documents with similarities outside the threshold. Support can be found in the specification on page 18, line 22-page 20, line 3. No new matter has been entered.

Such limitations are neither taught nor suggested by Lindh. Instead, Lindh teaches pre-processing that involves extract all terms included in unformatted text

10   and assigning weights to the terms based on their information content (*Lindh*, col. 16, lines 5-8). The information content can be determined by using an extension to a traditional Term Frequency times Inverse Document Frequency (TFIDF) term weighting scheme (*Lindh*, col. 17, line 20-col. 18, line 14). Lindh further teaches one method for enhancing relationship quality by filtering the document corpus

15   used to generate a term-term matrix (*Lindh*, col. 27, lines 18-25). The reduction in the number of similar documents in the corpus can result in large quantities of similar documents not biasing the relationship measures, which is characterized as a flaw that can be reduced using document clustering, such as *k*-means clustering (*Lindh*, col. 27, line 25-col. 28, line 5). A representative document vector is

20   generated for each cluster found by means of a clustering algorithm, such as by calculating a cluster centroid as the mean of all document vectors in the cluster (*Lindh*, col. 28, lines 8-23). A representative document vector is added to the cluster and all other documents that belong to the cluster are removed from the initial document corpus (*Id.*).

25   However, Lindh fails to teach or suggest a threshold module dynamically determining a threshold for each cluster based on similarities between the documents grouped into the cluster and a center of the cluster, and reassigning those documents having similarities outside the threshold, per Claim 1. Lindh also fails to teach or suggest dynamically determining a threshold for each cluster

30   based on similarities between the documents grouped into the cluster and a center

- 21 -

of the cluster; and reassigning those documents having similarities outside the threshold, per Claim 9. Lindh also fails to teach or suggest code for dynamically determining a threshold for each cluster based on similarities between the documents grouped into the cluster and a center of the cluster; and code for

5 reassigning those documents having similarities outside the threshold, per Claim 17. Lindh also fails to teach or suggest a threshold module dynamically determining a threshold for each cluster based on similarities between the documents in each cluster and the cluster center, and reassigning the documents with similarities outside the threshold, per Claim 18.

10 Lindh fails to disclose all of the claim limitations and does not anticipate Claims 1, 9, 17, and 18. Claims 2-8 are dependent upon Claim 1 and are patentable for the above-stated reasons for independent Claim 1, and as further distinguished by the limitations recited therein. Claims 10-16 are dependent upon Claim 9 and are patentable for the above-stated reasons for independent Claim 9,

15 and as further distinguished by the limitations recited therein.

As a *prima facie* case of anticipation has not been shown, withdrawal of the rejection for anticipation under 35 U.S.C. § 102(a) is requested.

Claims 24-27 and 41-47 stand rejected under 35 U.S.C. § 103(a) as being obvious over Lindh in view of U.S. Patent No. 6,675,159, to Lin et al. ("Lin").

20 Applicant traverses the rejection.

To establish a *prima facie* case of obviousness, the examiner must establish, *inter alia*, that the references teach the invention claimed. *In re Wood*, 599 F.2d 1032, 202 U.S.P.Q. 171, 174 (C.C.P.A. 1979). As Claims 24-27 are dependent upon Claim 18, these claims are patentable for the same reasons as

25 stated for Claim 18, and as further distinguished by the limitations recited therein. Similarly, as Claims 41-47 are dependent upon Claim 35, these claims are patentable for the same reasons as stated for Claim 35, and as further distinguished by the limitations recited therein.

As a *prima facie* case of obviousness has not been shown, withdrawal of

30 the rejection under 35 U.S.C. § 103(a) is requested.

- 22 -

Claims 28, 29, 32-35, 41-47, and 49-53 stand rejected under 35 U.S.C. § 103(a) as being obvious over Lindh in view of U.S. Patent No. 5,794,236, to Mehrle. ("Mehrle"). Applicant traverses the rejection.

To establish a *prima facie* case of obviousness, the examiner must
5   establish, *inter alia*, that the references teach the invention claimed. *In re Wood*, 599 F.2d 1032, 202 U.S.P.Q. 171, 174 (C.C.P.A. 1979).

Claim 35 has also been amended to further clarify the inventive subject matter and distinguish over the Lindh-Mehrle references. No new matter has been entered. Amended Claim 35 now recites scoring a document in an electronically-
10   stored document set and grouping the documents by score into a plurality of clusters. Support can be found in the specification on page 10, lines 17- 30. Claim 35 further recites identifying candidate seed documents, which are each assigned as a seed document into a cluster with a center most similar to the seed document; and assigning each non-seed document to the cluster with the best fit. Support can
15   be found in the specification on page 18, line 22-page 20, line 3. Claim 35 further recites dynamically determining a threshold for each cluster based on similarities between the documents in each cluster and the cluster center; and reassigning the documents with similarities outside the threshold. Support can be found in the specification on page 18, line 22-page 20, line 3. No new matter has been entered.

20   Claim 52 has been rewritten in independent form to define computer-readable storage mediums holding code with positive recitations of structural elements. Claim 52 similarly distinguishes over Lindh. Claim 52 recites code for scoring a document in an electronically-stored document set and code for grouping the documents by score into a plurality of clusters. Support can be found
25   in the specification on page 9, lines 17-24 and page 10, lines 17- 30. Claim 52 further recites code for identifying candidate seed documents, which are each assigned as a seed document into a cluster with a center most similar to the seed document; and code for assigning each non-seed document to the cluster with the best fit. Support can be found in the specification on page 9, lines 17-24 and page
30   18, line 22-page 20, line 3. Claim 52 further recites code for dynamically

- 23 -

determining a threshold for each cluster based on similarities between the

documents in each cluster and the cluster center; and code for reassigning the

documents with similarities outside the threshold. Support can be found in the

specification on page 9, lines 17-24 and page 18, line 22-page 20, line 3.

5   Claim 53 has also been amended to further clarify the inventive subject

matter and distinguish over the Lindh-Mehrle references. No new matter has been

entered. Amended Claim 53 now recites means for scoring a document in an

electronically-stored document set and means for grouping the documents by

score into a plurality of clusters. Support can be found in the specification on page

10   8, line 26-page 9, line 11 and page 10, lines 17- 30. Claim 53 further recites

means for identifying candidate seed documents, which are each assigned as a

seed document into a cluster with a center most similar to the seed document; and

means for assigning each non-seed document to the cluster with the best fit.

Support can be found in the specification on page 8, line 26-page 9, line 11 and

15   page 18, line 22-page 20, line 3. Claim 53 further recites means for dynamically

determining a threshold for each cluster based on similarities between the

documents in each cluster and the cluster center; and means for reassigning the

documents with similarities outside the threshold. Support can be found in the

specification on page 18, line 22-page 20, line 3. No new matter has been entered.

20   Such limitations are neither taught nor suggested by the Lindh-Mehrle

combination. The teachings of Lindh are discussed above with reference to the

rejection under 35 U.S.C. § 102(a). Instead, Mehrle teaches a system for

classifying documents into a hierarchy and linking the classifications into the

hierarchy (Mehrle, col. 2, lines 31-40). A scheme of unique classification keys are

25   mapped to each classification in the hierarchy to allow cumulative retrieval of all

classifications occurring below any non-terminal node of the hierarchy (Mehrle,

col. 6, lines 28-43). After all seed citations have been checked against seed

citations, citations scores are checked against a threshold value, which is set based

on the desired accuracy level of the system (Mehrle, col. 9, lines 39-53 and col.

30   10, lines 4-5). The accuracy level can be predetermined by the builder and a

- 24 -

threshold value for each classification can be entered based on the desired accuracy level. If a classification score is equal to or exceeds the threshold, the document receives the classification associated with the classification score (*Mehrle*, col. 9, lines 6-15).

5 However, the Lindh-Mehrle combination fails to teach or suggest *dynamically* determining a threshold for each cluster based on similarities between the documents in each cluster and the cluster center, and reassigning the documents with similarities outside the threshold, per Claim 18 (emphasis added). Rather, the threshold taught by Lindh-Mehrle is <u>static</u> and predetermined by the

10 builder of the system, instead of being dynamically determined. The Lindh-Mehrle combination also fails to teach or suggest code for *dynamically* determining a threshold for each cluster based on similarities between the documents in each cluster and the cluster center, and code for reassigning the documents with similarities outside the threshold, per Claim 52 (emphasis added).

15 Rather, the threshold taught by Lindh-Mehrle is <u>static</u> and predetermined by the builder of the system, instead of being dynamically determined by code. The Lindh-Mehrle combination also fails to teach or suggest means for *dynamically* determining a threshold for each cluster based on similarities between the documents in each cluster and the cluster center, and means for reassigning the

20 documents with similarities outside the threshold, per Claim 53 (emphasis added). Rather, the threshold taught by Lindh-Mehrle is <u>static</u> and predetermined by the builder of the system, instead of being dynamically determined by means.

As Claims 28 and 29 are dependent upon Claim 18, these claims are patentable for the same reasons as stated for Claim 18, and as further

25 distinguished by the limitations recited therein. Similarly, as Claims 41-44, 46, and 47 are dependent upon Claim 35, these claims are patentable for the same reasons as stated for Claim 35, and as further distinguished by the limitations recited therein.

As a *prima facie* case of obviousness has not been shown, withdrawal of

30 the rejection under 35 U.S.C. § 103(a) is requested.

- 25 -

Claim 30 stands rejected under 35 U.S.C. § 103(a) as being obvious over Lindh and Mehrle in view of Lin. Applicant traverses the rejection.

To establish a *prima facie* case of obviousness, the examiner must establish, *inter alia*, that the references teach the invention claimed. *In re Wood*,

5   599 F.2d 1032, 202 U.S.P.Q. 171, 174 (C.C.P.A. 1979). As Claim 30 is dependent upon Claim 18, this claim is patentable for the same reasons as stated for Claim 18, and as further distinguished by the limitations recited therein.
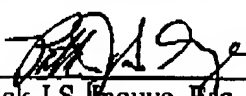
As a *prima facie* case of obviousness has not been shown, withdrawal of the rejection under 35 U.S.C. § 103(a) is requested.

10   The prior art made of record and not relied upon has been reviewed by the applicant and is considered to be no more pertinent than the prior art references already applied.

Claims 1-31, 35-44, 46-48, 52, and 53 are believed to be in condition for allowance. Payment for Claims 17 and 52, which have been rewritten in

15   independent form, is included. Entry of the foregoing amendments is requested and a Notice of Allowance is earnestly solicited. Please contact the undersigned at (206) 381-3900 regarding any questions or concerns associated with the present matter.

Respectfully submitted,

20
Dated: September 18, 2006   By: _____
                                        Patrick J.S. Inouye, Esq.
                                        Reg. No. 40,297

25   Law Offices of Patrick J.S. Inouye
     810 Third Avenue, Suite 258   Telephone: (206) 381-3900
     Seattle, WA 98104   Facsimile: (206) 381-3999

     OA Resp

- 26 -